13. Gambus, A., van Deursen, F., Polychronopoulos, D., Foltman, M., Jones, R.C., Edmondson, R.D., Calzada, A., and Labib, K. (2009). A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. EMBO J *28*, 2992–3004.

14. Pachlopnik Schmid, J., Lemoine, R., Nehme, N., Cormier-Daire, V., Revy, P., Debeurme, F., Debre, M., Nitschke, P., Bole-Feysot, C., Legeai-Mallet, L., *et al.* (2012). Polymerase epsilon1 mutation in a human syndrome with facial dysmorphism, immunodeficiency, livedo, and short stature ("FILS syndrome"). J. Exp. Med. *209*, 2323–2330.

Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, CB2 1QN, UK.
E-mail: paz20@hermes.cam.ac.uk

http://dx.doi.org/10.1016/j.cub.2013.03.008

Quantitative Genetics: Heritability Is Not Always Missing

Quantitative trait loci (QTLs) underlying multifactorial disorders explain little of the heritability — most is 'missing'. A new yeast study has identified QTLs which explain most heritability in traits. Why is heritability missing in human diseases but not here?

John F.Y. Brookfield

Human genetics once focussed on genes mutated in single-gene disorders. Now, however, human geneticists study the much more frequent multifactorial diseases, which result from the effects of many genes and from environmental effects. Having a genetic variant can raise the probability of the individual developing the condition, with the effect of a variant being quantified by the relative risk — the probability of the disease in someone carrying the genetic marker, relative to the probability in the general population. The identification of many such markers allows the prediction of an individual's genetic risk. This is potentially of benefit, as lifestyle changes or drug treatment can lower the risk even in the presence of genetic 'risk alleles'. The method used to identify causative quantitative trait loci (QTLs) is called a 'genome-wide association study' (GWAS) [1]. In GWASs, hundreds of thousands of single-nucleotide polymorphisms (SNPs) are compared between affected individuals and matched controls. Variants that show differences in frequency between the two groups are identified, with variants that are more common among affected individuals being inferred to create a relative risk above one. For a large number of diseases, such as type II diabetes, Crohn's disease and many other conditions, predisposing variants have indeed been found in this way [2,3]. However, at the same time, the heritability of the trait (in the narrow sense: the proportion of the observed phenotypic variance that is due to the additive genetic variance) can be estimated from correlations between relatives, and a paradox emerges: 'missing heritability'. When the effect sizes and frequencies of the known causative SNPs are combined they together account for only a small fraction of the overall heritability of the condition. For example, the 32 loci identified that contribute to risk of Crohn's disease only explain 20% of the heritability of the disease [3]. The same is seen in non-pathogenic variation; for instance, more than 50 loci have been identified as affecting human height, but collectively they account for only 5% of the narrow sense heritability, of around 80%, in height [4–8]. In trying to understand potential explanations of missing heritability, a recent study by Bloom, Kruglyak and colleagues [9] examining multiple traits in budding yeast (*Saccharomyces cerevisiae*) has demonstrated an example where there is very little missing heritability. Instead, the QTLs identified account for almost all the measured heritability. What are the implications of this finding for human quantitative traits?

Where Is the Missing Heritability?

There are many possible explanations for missing heritability. A technical one stems from the problem of how causative QTLs are ascertained. With hundreds of thousands of SNPs tested for disease associations, the threshold for statistical significance is high and is typically chosen on the basis of ensuring a low 'false discovery rate', of perhaps 5%. In other words, only 5% of the SNPs identified as QTLs for the disease will be false positives. This represents a much higher threshold for significance than would be adopted if only a single candidate locus were to be tested. Thus, for reasonable sample sizes, only QTLs with large effects can be identified with statistical confidence. It could thus be possible that very many SNPs have effects on the trait that are too small to be demonstrated, but which collectively could account for much or all of the missing heritability. There is indeed some evidence for this: Yang *et al.* [10] examined variation in height in an Australian population and found that the 294,000 SNPs examined could, when fitted collectively, explain 45% of the phenotypic variance, suggesting that there are many loci with additive effects on the trait in addition to those that reach a threshold for significance. This experiment also illustrates that, through use of genome-wide SNP information, more precise heritability estimation is also possible [11]. Heritability is estimated through phenotypic correlations between relatives. Thus, for example, full siblings should, on average, share half their DNA, but individual sibling pairs, through the random recombination and segregation events in their parents' germ cells, will share slightly less or more than this expected value. These differences can be observed by the use of SNP markers and make it possible, for example, to estimate heritability by correlating phenotypic similarity between sibling pairs with the proportion of the genome that they share.

Park *et al.* [12] examined the effect sizes detected with confidence in GWAS studies, and tried to estimate the underlying distribution of effect sizes. Thus, if a variant has a small effect, its probability of being detected in a given sample size can be calculated and thus the number of non-ascertained loci with equivalent effect sizes can be estimated. This allows the prediction of the extra causative loci that would be detected, and the extra heritability that would be

explained, through the use of larger sample sizes.

While the missing heritability may arise from common alleles of small effect, it could also be due to alleles that are found only very rarely in the population. These could be of large effect and yet still not be detected by GWAS. Most neutral variants will be rare, and disease variants will often be subject to purifying natural selection, further lowering their frequency [13]. Indeed, studies of very rare SNPs detected by large-scale human exon sequencing reveal that they are more likely to cause amino acid changes in the proteins they encode, relative to more abundant SNPs; this pattern is consistent with the action of purifying natural selection [14]. There are also other possibilities [4]: structural variants could contribute to heritability while being under-reported by SNP studies. For example, a rare 3 megabase deletion in chromosome 22 explains 1% of schizophrenia cases [15]. There could also be trans-generational non-genetic effects, creating correlations between relatives that augment the estimated heritability.

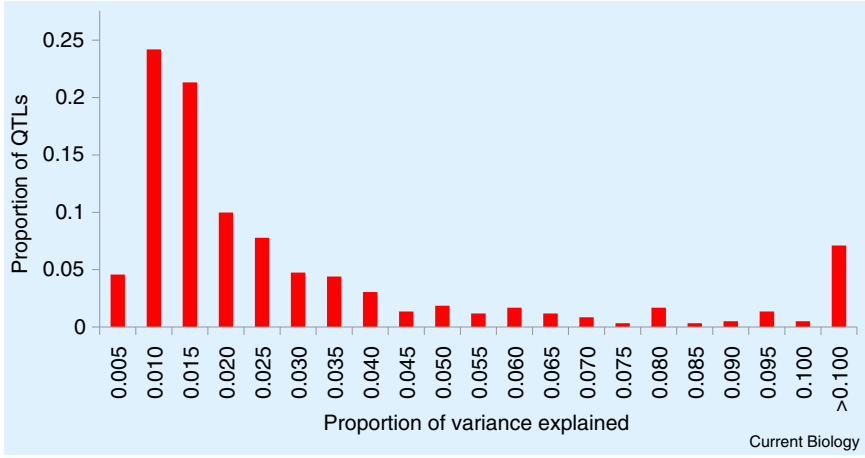

Figure 1. The distribution of effect sizes of QTLs in the yeast cross.
The graph shows the distribution of effect sizes (defined as the proportion of phenotypic variance explained) for the 591 yeast QTLs discovered by Bloom *et al*. [9]. The 27 loci affecting human height discovered by Gudbjartsson *et al*. [8] all have effect sizes below 0.005.

Missing Missing Heritability

The new study by Bloom *et al.* [9] did not look at human variation, but at a cross between two yeast strains (*S. cerevisiae*). The data were derived from 1008 haploid segregants from a sexual cross. While, on average, these haploid lines share half their genes, the SNP measurements on the segregant lines demonstrated the expected random variation in the extent to which their genomes were shared [11]. From this, heritability was estimated for each of 46 traits, all defined as growth rates under different environmental conditions. The median narrow sense heritability of the traits was 52%. While all the traits were growth rates, the correlations across strains in pairs of traits were typically low. For example, the Spearman rank correlation between growth on paraquat and growth on cobalt chloride was 1%. From the data, it was possible to identify causative QTLs that were supported, given a false discovery rate of five percent. In this way, 591 QTLs were discovered to be influencing the 46 traits, with between 5 and 29 QTLs per trait. But most notable here, there was almost no missing heritability at all. The detected QTLs explained between 72% and 100% (for the different traits) of the measured narrow sense heritability, with a median of 88%.

But what does this observation mean for missing heritability in human diseases? A yeast cross is very different from a human population. Most notably, as the genotypes tested come from segregants in a sexual cross, all variable SNPs will be at frequencies of almost exactly 50% in the 1008 strains tested by Bloom *et al.* [9]. The data suggest that, for these traits in yeast, the strain differences do not include large numbers of variants with very small effects on the traits, although some variants of small effect have probably not been ascertained (Figure 1). Compared to the results from human GWAS studies, there are fewer QTLs detected, yet these explain more of the heritability. Of course, the two strains will sample only a small subset of the variants that are present in the species, and will differ in variants that have both high and low minor allele frequencies in the yeast population. The question is whether this result can be extrapolated from the yeast situation to the human one, and whether the yeast growth traits examined are good proxies for human disease traits. The yeast strains differed in their DNA sequences by 0.5%, far more than the difference between any two human genomes, and one was a wine strain, which could have been subject to artificial selection.

In addition, the yeast genome is small, and thus will be a smaller mutational target. It is not clear what would be the expected relationship between genome size, or even gene number, and the distribution of effect sizes of new mutations that affect a specific trait. Without knowing this, we cannot predict whether an expected number of segregating functional variants will scale linearly with genome size. And it is unclear what would have been the past selective forces acting on the polymorphisms at which the two yeast strains differ, and the extent to which the traits measured would be correlated with fitness in the yeast's ancestry. If these traits have been subject to strong selection, the distribution of allele frequencies at the causative loci will differ systematically from that expected were the trait variation to be neutral.

In general, the understanding of the causes of the genetic variation affecting any trait, in any population, be it human or yeast, must come from understanding the population genetic history. Mutation, genetic drift and selection in the past have combined to create the standing genetic variation and the genetic architecture of traits.

References

1. Flint, J. (2013). GWAS. Curr. Biol. *23*, R265–R266.
2. Zeggini, E., Scott, L.J., Saxena, R., and Voight, B.F.; 3–5,7,10, for the Diabetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat. Genet. *40*, 638–645.
3. Barrett, J.C., Hansoul, S., Nicolae, D.L., Cho, J.H., Duerr, R.H., Rioux, J.D., Brant, S.R., *et al*. (2008). Genome-wide association defines

more than 30 distinct susceptibility loci for Crohn's disease. Nat. Genet. *40*, 955–962.
4. Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., *et al*. (2009). Finding the missing heritability of complex diseases. Nature *461*, 747–753.
5. Visscher, P.M. (2008). Sizing up human height variation. Nat. Genet. *40*, 489–490.
6. Weedon, M.N., Lango, H., Lindgren, C.M., Wallace, C., Evans, D.M., Mangino, M., Freathy, R.M., *et al*. (2008). Genome-wide association analysis identifies 20 loci that influence adult height. Nat. Genet. *40*, 575–583.
7. Lettre, G., Jackson, A.U., Gieger, C., Schumacher, F.R., Berndt, S.I., Sanna, S., Eyheramendy, S., *et al*. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. Nat. Genet. *40*, 584–591.
8. Gudbjartsson, D.F., Walters, G.B., Thorleifsson, G., Stefansson, H., Halldorsson, B.V., Zusmanovich, P., Sulem, P., *et al*. (2008). Many sequence variants affecting diversity of adult human height. Nat. Genet. *40*, 609–615.
9. Bloom, J.S., Ehrenreich, I.M., Loo, W.T., Volite, T.-V., and Kruglyak, L. (2013). Finding the sources of missing heritability in a yeast cross. Nature *494*, 234–237.
10. Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A., Heath, A.C., *et al*. (2010). Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. *42*, 565–569.
11. Visscher, P.M., Medland, S.E., Ferreira, M.A.R., Morley, K.I., Zhu, G., Cornes, B.K., Montgomery, G.W., and Martin, N.G. (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. PLoS Genet. *2*, e41.
12. Park, J.-H., Wacholder, S., Gail, M.H., Peters, U., Jacobs, K.B., Chanock, S.J., and Chatterjee, N. (2010). Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat. Genet. *42*, 570–575.
13. Eyre-Walker, A. (2010). Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. Proc. Natl. Acad. Sci. USA *107* (*Suppl. 1*), 1752–1756.
14. Marth, G.T., Yu, F., Indap, A.R., Garimella, K., Gravel, S., Leong, W.F., Tyler-Smith, C., *et al*. (2011). The functional spectrum of low-frequency coding variation. Genome Biol. *12*, R84.
15. Bassett, A.S., Marshall, C.R., Lionel, A.C., Chow, E.W., and Scherer, S.W. (2008). copy number variations and risk for schizophrenia in 22q11.2 deletion syndrome. Hum. Mol. Genet. *17*, 4045–4053.

Centre for Genetics and Genomics, School of Biology, University of Nottingham, University Park, Nottingham NG7 2RD, UK.
E-mail: John.Brookfield@nottingham.ac.uk

http://dx.doi.org/10.1016/j.cub.2013.02.040

Animal Behavior: Fly Flight Moves Forward

A new study has resolved the paradox of how flies maintain reflexive aversion to your approaching swatter, whilst tolerating similar visual signals during normal forward flight.

Jessica L. Fox and Mark Frye

Animal nervous systems come equipped with many built-in rapid reflexes. These simple behaviors maintain an animal's physical stability without requiring much neural overhead, and they permit animals to respond rapidly in situations where taking the time to make complex calculations within the central nervous system would jeopardize survival. During rapid locomotion, reflexes are crucial to keeping the body in its correct posture and responding to external perturbations, such as an obstacle in the path of a runner or a gust of wind knocking a bird off of its flight path. Low-level sensory-motor reflexes enable the nervous system to efficiently maintain control while walking, swimming, or flying through varied and often unpredictable environments. Managing these reflexes, however, can be complicated: the nervous system must have means by which the reflexes can be deployed in the right circumstances, and suppressed in the wrong ones. A new study by Reiser and Dickinson [1] has revealed how the visual systems of flies in flight use a surprisingly simple algorithm to decide when a reflex should be employed, and when it should be overridden.

In insects, a reflex known as the looming avoidance response keeps the animals from colliding with objects, or becoming snatched by oncoming predators. The image of an object on the retina increases in size as the object gets closer, and the rate of change in image size will speed up as the object and the observer close in on one another. Insect nervous systems are able to use this rate of change in perceived size to calculate the time to collision, and standing insects will jump to avoid the oncoming object [2,3], whereas flies in flight similarly turn away from looming objects [4]. An approaching object generates optic flow across the eye that forms a 'vanishing point': the optic flow pattern expands outward along the direction of motion from the focus of expansion and disappears in a point behind the observer (the focus of contraction). Expanding optic flow on its own triggers collision-avoidance turning reflexes in flies [4], suggesting that the looming avoidance response is not selective for objects in particular, but is a more generalized reflex for avoiding any image expansion. Focus-of-expansion avoidance is obviously a useful response for preventing impact or dodging a swatter, but it presents the fly with a paradox: if it turns away from all looming signals, how can it ever fly forward?

Over the past five years, three papers from the Dickinson lab have found common scenarios in which the fly overrides its escape reflex to fly towards a focus of expansion instead of turning away from it. First, Budick *et al.* [5] found that flies will more readily orient towards a focus of expansion if they can fly against a gentle headwind, which would be induced by normal forward flight in still air. Also, noting that a fly will readily approach a vertically-oriented object representing a landing perch or a gap in the foliage, Reiser and Dickinson [6] found that placing such an object within the focus of expansion switched their behavior from expansion avoidance to object tracking, and thus permitted them to fly forward into the focus of expansion. Now, Reiser and Dickinson [1] have further demonstrated that the focus-of-expansion avoidance response is dependent on the strength or velocity of the expanding optic flow emanating from it, and that if the expansion velocity is sufficiently low, then flies will fly towards rather than away from the focus of expansion, even without oncoming wind or any other attractive feature.

Reiser and Dickinson [1] used an electronic flight simulator to present visual stimuli to tethered fruit flies in flight. In this arena, the flies can flap their wings, but cannot move their